

**DİLÇİLİK  
ЯЗЫКОЗНАНИЕ  
LINGUISTICS**

*АЙМАН ЖАНАБЕКОВА (Казахстан)\**  
*КУНСУЛУ ПИРМАНОВА (Казахстан)\*\**  
*САНСЫЗБАЙ МАДИЕВ ШАРАПИУЛЫ (Казахстан)\*\*\**  
*БУЛБУЛ КАРБОЗОВА (Казахстан)\*\*\*\**

**РАЗРАБОТКА ЛЕКСИКО-СЕМАНТИЧЕСКОЙ РАЗМЕТКИ В  
НАЦИОНАЛЬНОМ КОРПУСЕ КАЗАХСКОГО ЯЗЫКА****Резюме**

В статье дается обзор того, какие принципы изначально необходимы для разработки лексико-семантической разметки казахского национального корпуса. В целом речь идет о двух направлениях корпуса. Об аннотированном и неаннотированном корпусах.

Корпусная лингвистика поможет обучающимся казахскому языку, языковедам, лингвистам и другим пользователям корпуса овладеть теоретическими и практическими знаниями казахского языка. Что касается теории мировых языковых корпусов, то данных, касающихся внутренней подкорпусной функции корпуса, достаточно. В корпусе языка есть лингвистические классификации уровней языка. Решение практических задач по всем уровням языка очень важная работа при разработке лексико-семантических и других разметок. В данной статье описаны практические задачи для разработки лексико-семантической разметки. При составлении алгоритма для корпуса, касающегося данной разметки, мы опирались на модель частотности падежей казахского языка, которая была составлена основателем прикладного направления в казахском языкознании, профессором А.К. Жубановым.

В национальном корпусе казахского языка также есть трудности в разработке лексико-семантической разметки. Несмотря на то, что различение значения слова в языке, познание его смысла не составит трудностей для лингвистов, простому поль-

---

\* Институт языкознания имени А.Байтурсынова Министерства образования и науки Казахстана, доктор филологических наук. E-mail: aiman\_miras@mail.ru

\*\* Казахский национальный университет им. аль-Фараби Институт языкознания имени А.Байтурсынова Министерства образования и науки Казахстана, научный сотрудник.  
E-mail: kunsulu.pirmanova@mail.ru

\*\*\* Институт языкознания имени А.Байтурсынова Министерства образования и науки Казахстана, кандидат филологических наук. E-mail: sansyzbay.madiev@mail.ru

\*\*\*\* Институт языкознания имени А.Байтурсынова Министерства образования и науки Казахстана, доктор философии. E-mail: bulakarb\_83@mail.ru

зователю нелегко справиться с данной задачей, касающейся семантической стороны слова. Важным делом в овладении языком является разработка лексико-семантических разметок к каждому слову в корпусе языка.

**Ключевые слова:** *корпусная лингвистика, национальный корпус казахского языка, разметка, лексико-семантическая разметка, принципы семантической классификации лексики казахского языка, семантическая классификация имен существительных, семантическая классификация глаголов.*

При изучении казахского языка необходимо уделять большое внимание области корпусной лингвистики, ее теоретическим и практическим аспектам мирового уровня. В специальных изданиях научных журналов также публикуются статьи по общим и конкретным вопросам, касающимся создания и работы текстовых корпусов по всему миру.

Корпуса делятся на различные подкорпусы в зависимости от их функции. Создание аннотированных корпусов, особенно лингвистических разметок для лингвистических исследований, очень важно и полезно. В зависимости от того, поставлена эта разметка или нет, корпуса делятся на аннотированные и неаннотированные.

Типы разметок, которые прикреплены к корпусу, можно разделить на две категории: лингвистические и внешние. Лингвистические разметки включают в себя: лексико-семантическую, морфологическую, словообразовательную, фонетическую и другие разметки.

В национальном корпусе казахского языка существуют все вышеуказанные разметки. Однако в лексико-семантической разметке не хватает полного описания семантических групп слов. Для решения этого вопроса нам необходимо делить все казахские слова на лексико-семантические группы. В казахском языке исследователи по-разному делят слова на лексико-семантические группы. Далее будет представлена классификация профессора А.К.Жубанова.

В соотношении «язык – реальная действительность» язык служит не только обозначением реальной действительности, но в то же время является носителем обобщенного отражения действительности, органически соединяясь с соответствующими понятиями или отдельными их признаками, лежащими в основе семантических ценностей языковых единиц.

Общеизвестно, что лексические средства, наряду со звуковыми и грамматическими, составляют основу любого языка, в частности, и казахского. Слово, называя предметы, явления реального мира в их обобщенном виде, выступает в языке основным средством номинации. Являясь структурной ячейкой строя языка, оно имеет определенную морфологическую оформленность и грамматическую значимость. Слово охватывает фактически все языковые функции: номинативную (обозначения), сигнификативную (обобще-

ния), коммуникативную (общения) и прагматическую (экспрессивного выражения и эмоционального воздействия).

Словарные знаки слов таких классов, как существительные, глаголы и прилагательные, обладают двумя основными семиологическими функциями: наименования и описания. Совместимость их в пределах слова порождает назывные высказывания, предикативные и лексические синтагмы. Они являются тем основным звеном в словарном составе языка, той постоянной значимой основой, которая обеспечивает «переход» от языка к мысли, от системы к речи, от слов к синтагматическим единицам, большим, чем слова.

При семантической классификации лексики казахского языка имеются свои трудности и своя специфика. Это одна из труднейших задач семантики, которая не находит единого решения и в других языках мира. Количество таких классов и подклассов, их состав зависят от целей исследования, лингвистической школы и целого ряда других факторов.

Работ, посвященных исследованию казахского языка в области семантики, относительно мало. Здесь стоит назвать имя М.Оразова, его монографию «Қазақ тілінің семантикасы» и его исследования по семантике казахского глагола (Оразов, 1991; Оразов, 1983), которые затрагивают многие вопросы семантики казахского языка. Имеется и ряд других работ по некоторым проблемным вопросам семантики казахского языка (Ахметжанова, 1989). Хотелось только отметить, что результаты этих работ во многом помогли нам при решении поставленных перед нами задач по вопросам формализации содержания казахского текста.

В самом общем виде можно выделить три принципа семантической классификации лексики казахского языка:

- 1) денотативно-десигнативный;
- 2) парадигматический;
- 3) синтагматический (Оразов, 1991; Васильев, 1981; Zubov, 1985).

Денотативно-десигнативный принцип учитывает онтологическое расчленение действительности. При этом слова объединяются в группы в соответствии с тем, что они обозначают в реальной действительности (названия животных, названия растений, наименования определенных классов предметов и т.п.). При парадигматической классификации слова объединяются в группы на основе тождества определенных семантических компонентов, выделенных в значении слов (модальные и каузативные глаголы и т.п.). По синтагматическому принципу слова объединяются в лексико-семантические группы с учетом количества и качества валентностей классифицируемых слов (Оразов, 1991: 189-192; Васильев, 1981: 39-42).

Исходя из целей исследования за основу нашей семантической классификации принят денотативно-десигнативный принцип, хотя в некоторых случаях он будет дополняться двумя другими перечисленными выше принципами.

**Семантическая классификация имен существительных (зат есім) казахского языка.** Как известно, основным критерием разграничения имен существительных (наименований явлений, предметов, процессов, понятий) казахского языка, как и в других языках, на отдельные типы является прежде всего анализ их прямого номинативного значения, детерминированного типом денотата и сигнификата. При этом «слово» можно рассматривать двояко, т.е. как единицу первичного означивания в системе номинативных средств языка и как единицу, семантически реализованную в речи. Нами различаются две модификации «денотатов»: денотат как объект обозначения единичного предмета или класса предметов, а также денотат как предмет референции знака в речи.

А.А.Уфимцева в своей книге «Типы словесных знаков» (Уфимцева, 1974) подразделяет сигнификат как компонент знакового значения слов на шесть понятий, которые, как нам кажется, справедливы и для казахского языка. Имена с денотативным и денотативно-сигнификативным значением относятся к разряду имен, называемых конкретной лексикой, а слова с сигнификативной и сигнификативно-денотативной основой значения составляют абстрактную лексику любого языка, в том числе и казахского. Необходимо заметить, что определяющим для семантики именных лексем является обязательное наличие в их знаковом значении денотата и сигнификата. Если словесный знак служит специализированным названием (термином) в какой-либо узкой предметной области, то доминирует денотат над сигнификатом. В случае, когда словесный знак служит названием абстрактного понятия, то, наоборот, или сигнификат доминирует над денотатом, или словесный знак имеет чисто сигнификативное значение.

Именами существительными в казахском языке, так же, как и во многих тюркских языках, являются различные наименования предметов и понятий в широком смысле слова. Это различные имена: собственные, например, имена людей, географические названия и т.д.; нарицательные – названия человека: *кісі* (человек), *шеше* (мать), *әйел* (женщина); названия животных; названия растений; названия различных веществ и предметов быта; научные, технические и другие термины; наименования разных обобщенных отвлеченных понятий: *ел* (народ), *ру* (род), *тағдыр* (судьба) и т.д.; наименования действий, явлений, состояний: *жүріс* (ходьба), *хабар* (известие), *ұйқы* (сон) и т.д.; названия результатов действия: *жиналыс* (собрание; ср. *жинау* – собираться), *кеспе* (лапша; ср. *кесу* – резать); названия орудий (инструментов) действия: *бұрғы* (сверло; ср. *бұрау* – винтить, завинчивать), *сытырғыш* (метла, веник), *ара* (пила) и т.п.

Как пишет М.Оразов, среди слов казахских существительных самыми частыми по сложности их значений являются нарицательные имена. Эти имена существительные по отношению к значению имеют в языке изменчивый характер, т.е. иногда у них превалирует денотативное значение, иногда сигнификативное или же приобретает как бы суммарное значение – и дено-

тативное, и сигнификативное. Примером может служить использование слов в метафорическом и метонимическом значениях. Безусловно, когда слова употребляются в переносном значении, то изменяется и их денотативное и сигнификативное значение. Для убедительности сказанного можно сравнить слово «нар» в двух следующих предложениях: *Қырық балам бар еді, қырқы да бірдей нар еді.* (Батырлар жыры). *Айырдан туған жампоз бар – Нарға жүгін салғызбас* (Махамбет) (Оразов, 1991: 98). В первом предложении слово «нар» употреблено в переносном значении как сильный, «боевой», а во втором – в прямом значении как одногорбный верблюд. Их денотативные и сигнификативные значения несравнимы.

В количественном отношении тематические группы имен существительных в любом языке, в том числе и в казахском языке, довольно велики. М.Оразов для примера приводит 8 разновидностей таких групп:

- 1) родственные термины (туысқандық атаулар);
- 2) человек и названия частей его тела (дене мүше атаулары);
- 3) названия домашних животных (төрт түлік мал атаулары);
- 4) названия одежды, обуви человека (киім-кешек атаулары);
- 5) название птиц (құс атаулары);
- 6) название диких животных (жабайы аң атаулары);
- 7) название обиходных предметов-сосудов (посуды) (ыдыс-аяқ атаулары);
- 8) названия болезни человека (адамға тән ауру атаулары).

В нашем исследовании в основу семантической классификации казахских существительных положено деление реальной действительности на определенные зоны, сегменты (Зубов, 1985 и др.). С учетом целей нашего исследования, анализируемого материала, а также отдельных идей, изложенных в работах Оразова, 1991; Ахметжановой, 1989; Апресяна, 1967, нами выделены семантические классы существительных (см. *табл. 1*). Как видно из таблицы, денотативно-десигнативный принцип использован здесь при выделении наиболее крупных семантических классов существительных. Более мелкие лексико-семантические группы, например, «части тела человека», «воинские звания», «дни недели» и др., связаны парадигматическими отношениями.

**Семантическая классификация глаголов. Знаковое значение глагольных лексем.** Имена существительные отражают и обозначают предметы или явления, а глаголы – отношения последних. Значение имен существительных – более автономно и абсолютно. Значение глагольных лексем – относительно и детерминировано многими факторами. В глагольных лексемах, выражающих понятие отношения, их взаимодействие как бы выносится за рамки глагольной лексемы и переносится в область синтагматических отношений действия к его объекту или субъекту (или одновременно). Поэтому соотносительность глагольных лексем с предметным рядом предполагает наличие смысловых связей глагола с предметными именами.

**Таблица 1. Семантические классы существительных,  
использованных в модели порождения казахского текста**

Сегменты реальной действитель- ности	Наименование семантического класса	Код класса
1. Наука	Общефилософские понятия	ЗТ01
	Общенаучные понятия	ЗТ02
	Общемедицинские понятия	ЗТ03
	Основные понятия психологии	ЗТ04
	Основные понятия нейропсихологии	ЗТ05
	Основные понятия логики	ЗТ06
	Основные понятия геометрии	ЗТ07
	Основные географические понятия	ЗТ08
	Основные понятия языкознания	ЗТ09
	Основные понятия литературоведения	ЗТ10
	Общемедицинские понятия	ЗТ11
	Основные понятия музыкального искусства	ЗТ12
	Основные юридические и уголовные понятия	ЗТ13
	Основные понятия логики	ЗТ14
	Основные понятия геометрии	ЗТ15
	Основные географические понятия	ЗТ16
	Основные понятия языкознания	ЗТ17
	Основные понятия литературоведения	ЗТ18
	Основные понятия изобразительного искусства	ЗТ19
	Основные понятия музыкального искусства	ЗТ20
	Основные юридические и уголовные понятия	ЗТ21
	Понятия, связанные с авиацией	ЗТ22
	Понятия, связанные с устройством отдельн. механ-го объекта	ЗТ23
	Информационные понятия	ЗТ24
2. Природа	Оценка	ЗТ31
	Предметы природы	ЗТ31
	Явления природы	ЗТ32
	Предметы природы – артефакты	ЗТ33
	Состояния природы	ЗТ34
	Понятия, описывающие структуру предм-в и части предметов	ЗТ35
	Вещества, металлы	ЗТ38
	Действия предметов природы	ЗТ39
3. Человек и его мир	Человек и части его тела	ЗТ40
	Части лица и головы человека	ЗТ41
	Состояния и ощущения, специфичные для человека	ЗТ42
	Действия, специфичные для человека	ЗТ43

	Понятия, используемые при характеристике человека	3Т44
	Предметы одежды и обуви чел-ка и понятия, связан. с этим	3Т45

	Предметы-артефакты, используемые человеком в целях гигиены, для украшения и т.п.	3Т46
	Объекты, используемые для передвижения человека, и понятия, с этим связанные	3Т47
	Предметы питания человека	3Т48
	Понятия, связанные с жилищем человека	3Т49
	Понятия, связанные со смертью и захоронением	3Т50
	Имена, отчества и фамилии людей	3Т51
	Национальность человека	3Т52
	Предметы-сосуды, которыми пользуется человек, и связанные с этим понятия	3Т53
	Предметы, используемые на отдыхе, охоте, рыбалке, и связанные с этим понятия	3Т54
	Наименование гражданских специальностей, должностей, степеней, званий	3Т55
	Предметы-артефакты, используемые при учении, труде и т.п.	3Т56
	Понятия, связанные с взаимоотношением людей	3Т57
	Понятия, связан. с родственными отношениями между людьми	3Т58
4. Животный мир	Награды человека	3Т59
	Крупные домашние животные и части их тела	3Т60
	Насекомые и части их тела	3Т61
	Птицы и части их тела	3Т63
	Действия, специфичные для животных	3Т64
	Предметы питания животных, птиц и насекомых	3Т65
5. Растения	Растения и их части	3Т66
	Деревья и их части	3Т67
6. Война и понятия, с ней связанные	Предметы боевой техники, оружия и их части	3Т68
	Наименование военных специальностей	3Т69
	Воинские звания	3Т70
	Понятия, связанные с оказанием помощи во время войны	3Т71
	Названия воинских подразделений	3Т72
	Понятия, связанные с войной	3Т73
7. Время	Временные понятия	3Т74
	Дни недели	3Т75
8. Религия	Понятия, связанные с обучением и трудом	3Т76
9. Учеба и труд	Организация учебы, труда	3Т77
	Понятия, связанные с обучением и трудом	3Т78
10. Меры	Меры длины, веса, скорости и т.п.	3Т79

11. География	Названия населенных пунктов	ЗТ80
	Названия стран, материков, частей света, островов, гор и др.	ЗТ81
	Названия рек, морей, озер, проливов	ЗТ82

Знаковое значение глагольных лексем представляет собой сложный феномен, детерминируемый тремя факторами: 1) соотносительностью с реальной действительностью, которая представляет собой в данном случае не мир предметов, а мир их отношений, действий и состояний; 2) категориальной семантикой сочетающихся с глаголом предметных имен; 3) типом смысловых отношений между действием, его субъектом и объектом, а соответственно типом моделей «объект – действие», «действие – объект», «субъект – действие – объект» (Уфимцева, 1974: 113).

По сравнению с именами существительными класс глагольных лексем представляет большие трудности при выделении типов по характеру их знакового значения.

Во-первых, собственно лексическое (индивидуальное) значение глагольных слов осложнено целым рядом семантических признаков, выражающих способ протекания действия, обстоятельства его свершения, видовременные характеристики глагольного действия, модальные и экспрессивные наслонения.

Во-вторых, определение глагольной семантики затруднено уточнением, какой из признаков присущ только данной лексеме, какой – целой группировке, а какой обобщен до уровня категории глагольных лексем и т.п.

Традиционно глаголы по их семантике делятся на три класса:

- 1) глаголы действия;
- 2) глаголы состояния;
- 3) глаголы процессов.

У каждой лексико-семантической группы – свои парадигматические параметры, семантические признаки. Так, для глаголов «движения», довольно хорошо изученных во многих языках, выделяются следующие семантические (лексические) категории: направленные движения, скорость движения, манера передвижения, средство или среда передвижения и т.п.

Для семантики эмотивных глаголов существует иная шкала – глаголы воздействия, глаголы восприятия, глаголы оценки или субъективного отношения и т.п. (Уфимцева, 1974: 148).

М. Оразов в своей книге «Қазақ тілінің семантикасы» пишет, что неконкретность лексического значения глагола в основном зависит от его денотативного и сигнификативного значений. Когда речь идет о денотативном значении глагола, мы говорим о движении предмета, физическом действии (амал-эрекет) и о состоянии (қалпы). Если учесть, что движение происходит в какое-то определенное время и в определенном месте, то имеется в виду присутствие субъекта, который осуществляет физическое действие, известна также цель передвижения. При этом лексическое значение глагола дополня-



ют множество других элементов. В качестве примера он приводит следующее предложение: *Шаңғышылар селого кіргенде көз байланып та қалған еді.* (М.Сундетов). */Когда лыжники вошли в село, уже наступили сумерки/*. В этом предложении рассматривается структура значения глагола *кір* (входи):

- 1) наличие движения;
- 2) движется некоторый предмет или человек (субъект);
- 3) движение происходит на земле;
- 4) переход через некоторую границу (вошли на территорию села);
- 5) имеется элемент значения войти «внутрь». Перечисленные значения

определяются благодаря отношениям и связям слов в данном предложении. Таким образом, действие *кіру* (входить) происходит на лыжах (на снегу) движением лыжников, *кіру* означает вхождение на территорию, где проживают люди (в селе). Благодаря сочетанию слова *кір* с другими словами (*село, шаңғышылар*), мы узнаем направление движения, что оно происходит «извне в село».

Глагол *кір*, сочетаясь с другими словами, может принять абсолютно другое значение. Например, *Мен 1962 жылы партия ұйымына кірдім*. В этом предложении слово *кір* уже не означает движение, а дает понять, что он стал членом партийной организации. Последний пример позволяет говорить о том, что значение слова непосредственно не зависит от значения словосочетаний (от контекста, в широком смысле) в предложении. В первом примере глагол *кір* сочетается с объектом, означающим местность, поэтому его граница является основой понимания лексического значения из сем движения. Во втором примере глагол *кір* сочетается с понятием, т.е. названием организации, здесь сема глагола *кір* единственная – стать членом данной организации. Изменение значения некоторых глаголов требует очень тонкого подхода при анализе их семантики. Например, глаголы движения в одних сочетаниях входят в структуру глагола общения, а в других – в структуру глагола мышления, Сравните: *үйге келу* (прийти домой), *қарсы келу* (возразить), *ойға келу* (вспомнить), *тілге келу* (заговорить) и др. В структуре предложения глаголы образуют отношения с определенными словами. Такие отношения происходят между субъектом или объектом и движением, между физическим действием (*амал-әрекет*) и объектом или субъектом. Когда в предложении глагол требует связи с другим словом, то он с этим словом имеет связь в отношении и значения, и формы (валентность глагола).

**Тематические и лексико-семантические классы казахских глаголов.** Основанием для отнесения слов к различным тематическим и лексико-семантическим классам являются обобщающие (жалпылаушы) семы. Необходимо отметить, что в индоевропейских языках имеется определенный опыт в разделении на лексико-семантические классы отдельных частей речи, но в отношении тюркских языков, в частности казахского, подобные исследования начали появляться только в последнее 10-летие (большинство из них имеют сопоставительный с русским языком характер). В настоящее время

описаны следующие лексико-семантические группы глаголов казахского языка: глаголы физиологического воздействия (Касабекова, 1985), глаголы речи (Курамысова, 1986), глаголы ориентированного положения в пространстве (Дуйсекова, 1996), глаголы движения (Шайкенова, 1981), эмотивные глаголы (Ибадильдина, 1997), глаголы физического восприятия (Уразымбетова, 1998), глаголы созидания (Токтарова, 1998) и др.

В своей семантической классификации знаменательных классов слов казахского языка старались придерживаться результатов исследований М.Оразова, а именно основных условий объединения слов в одну лексико-семантическую группу, а также и других работ вышеупомянутых ученых.

В центре лингвистической проблематики находится изучение фундаментальных семантических категорий со сложной формально-содержательной структурой. К числу таких категорий, отражающих одну из форм всеобщей связи явлений и фактов объективной действительности, следует отнести категорию каузальности (причинности).

В исследованиях казахского языка категория каузальности не была объектом специального изучения. Каузальные отношения в языке рассматривались в пределах отдельно взятых лингвистических задач. Например, в работах А.К.Калыбаевой-Хасеновой, С.В.Шкурдина и др. на морфемном уровне рассматривались глаголы в форме понудительного залога, образующие морфологическую каузативность. На синтаксическом уровне изучались сложно-подчиненные и бессоюзные сложные предложения с причинно-следственным характером отношений между составляющими их частями (М.Б.Балакаев, Т.Кордабаев, Н.Х.Демесинова, К.Есенов, Т.Ж.Тажибаева, У.Салиева и др.).

В последнее 10-летие научная школа, возглавляемая доктором филологических наук, профессором З.К.Ахметжановой, намного продвинула вперед функционально-семантическое направление исследования в сопоставительном плане изучения казахского и русского языков. Необходимо отметить, что при семантической классификации каузативных глаголов казахского языка мы в своей работе в основном придерживались положений работы Г.М.Алиной «Каузальные отношения в системе русского и казахского языков» (рукопись канд. дис., руков. З.К.Ахметжанова).

Необходимо отметить, что бытийные глаголы казахского языка еще не освоены до конца. В существующих классификациях (И.Маманова, А.Ыскакова, Н.Оралбаевой, М.Оразова, Н.Душимовой и др.) лексико-семантических групп казахского языка бытийные глаголы не выделяются как отдельная группа. Глагол «быть» является основным средством выражения понятия бытия в русском языке. В казахском языке глаголы сугубо бытийного значения, такие, как *еді* (*екен*, *емес*), *бол*, характеризуются в основном как вспомогательные, которые самостоятельно не употребляются, но образуют описательные формы времен и наклонений основных глаголов.

А. Ыскаков, в свою очередь, определяет глаголы, образованные от корня *-e* (*еді*, *екен*, *емес*), как неполные слова, которые утратили свое лексиче-

ское значение либо полностью, либо наполовину и играют в предложении вспомогательную роль, придавая грамматическое значение отдельным словам (*бала еді, қырық-отыздар шамасында екен, қара-торы ала емес*).

Учитывая все высказывания исследователей о семантических классах казахских глаголов, мы для наглядности решили свести их в одну таблицу (*табл. 2*). По нашей классификации семантический класс казахских глаголов делится на 10 разновидностей: I. Глаголы существования; II. Глаголы физических действий; III. Глаголы движения; IV. Глаголы, обозначающие отношения между людьми, животными и неживыми объектами; V. Глаголы психической деятельности; VI. Глаголы физического состояния и нормы; VII. Глаголы психического состояния; VIII. Фазисные глаголы; IX. Модальные глаголы; X. Каузативные глаголы.

Каждый класс глагола может быть разделен еще и на семантические подклассы. В данной статье мы решили показать лишь фрагмент, охватив некоторые семантические классы и подклассы казахских глаголов. В таблице показана следующая информация: 1. В отдельной строке под римской нумерацией дается семантический класс глаголов; 2. В первом столбце таблицы указаны семантические подклассы каждого класса; 3. Во втором столбце – соответствующие примеры к каждому подклассу казахских глаголов; 4. В третьем столбце даны соответствующие коды каждого класса и подкласса, которые будут использованы при записи формул предложений, абзацев и текста.

**Таблица 2. Семантические классы казахских глаголов, используемых в модели порождения текста (фрагмент)**

Семантический подкласс глагола	Примеры	Код
1	2	3
<b>Семантический класс: I. Глаголы существования</b>		
1. Бытийные глаголы: <i>Бар</i>	<i>Менде жақсы оқуға мүмкіншілік бар, бірақ оған түк ынтам жоқ.</i>	ET101
<i>Жоқ</i>	<i>Оның жасы әлі кәмелетке толған жоқ.</i>	ET102
Бол (болу, болмау)	<i>Сөзіңді айтып бол.</i> <i>Кеше жиналыс болды</i>	ET103
Е (еді, екен, емес)	<i>Алматы Қазақстанның астанасы еді. Өзі бір қызық адам екен.</i> <i>Оның аты қара-торы емес.</i>	ET104
<b>Семантический класс: II. Глаголы физических действий</b>		
1. Глаголы физического изменения: создания и производства, соединения и деления и т.д.	<i>Жасау, соғу, тігу, тұрғызу, салу, бөлшектеу, екшеу, ыдырату</i> и т.д.	ET201
2. Глаголы активных физических действий: разрушения и порчи	<i>Қирату, бүлдіру, бұзу, сындыру, құлату</i> и т.д.	ET202
3. Глаголы эмоциональных	<i>Аузы ашылу, мәз болу, күлу,</i>	ET203

физических действий	<i>ыржиу, шапалақтау, секіру (қу-анғаннан)</i>	
4. Глаголы, употребляющиеся при приеме пищи, свойственные для человека и животного мира	<i>Ишу, жеу, ұрттау, жәту, асау, дәм тату, сораптау</i> и т.д.	ЕТ204
5. Глаголы прочих физических действий, свойственных человеку и животному миру	<i>Жүгіру, секіру, ырғу, құсу, таптау, жылжу, қыймылдау ....</i>	ЕТ205
6. Глаголы физических действий, описывающих неживую природу	<i>Күркіреу, сарқырау, арсылдау, сілкіну</i> и т.д.	ЕТ206
7. Глаголы прочих физических действий, свойственных неживой природе	<i>Ағу, бату (күн), құлау, қайнау, дірілдеу, ысу, суу, гүрсілдеу</i> и т.д.	ЕТ207
8. Глаголы, связанные со звуками окружающей среды, глаголы звуковыражения	<i>Тарсылдау, сатырлау, сарқырау, ызыңдау, гүрсілдеу, қоңқылдау</i> и т.д.	ЕТ208
9 Глаголы физических действий, свойственных только животному миру	<i>Ұлу, ұру, ырылдау, мөңку, мөңіреу, жорту, шабу</i> и т.д.	ЕТ209
<b>Семантический класс: III. Глаголы движения</b>		
1. Глаголы, обозначающие движения, происходящие на земле	<i>Жүру, жүгіру, еңбектеу, домалау, сырғанау</i> и т.д.	ЕТ301
2. Глаголы, обозначающие движения, происходящие на воде (в воде)	<i>Жүзу, сүңгу, малту, қалқу</i> и т.д.	ЕТ302
3 Глаголы, обозначающие движения, происходящие в воздухе	<i>Ұшу, самғау, қалықтау, шүйлігу, зымырау</i> и т.д.	ЕТ303
4. Глаголы движения, обозначающие направленное движение	<i>Өрлеу, ылдилау, келу, кету, жақындау, ұзау</i>	ЕТ304
5. Глаголы движения, обозначающие хаотичность движения	<i>Аралау, шарлау, кезу, қыдыру, атқу</i> и т.д.	ЕТ305
<b>Семантический класс: IV. Глаголы, обозначающие отношения между людьми, животными и неживой природой</b>		
1. Глаголы выравнивания	<i>Алу, қосу, теңестіру, салыстыру, бөлу</i> и т.д.	ЕТ401
2. Глаголы принадлежности	<i>Иемдену, меншіктеу, иелену, баурау</i> и т.д.	ЕТ402
3. Глаголы приветствия и прощания	<i>Амандасу, сәлемдесу, көрісу, қоштасу</i> и т.д.	ЕТ403
4. Глаголы почтения	<i>Құрметтеу, қадірлеу, ізеттеу, сыйлау</i> и т.д.	ЕТ404
5. Глаголы властвования	<i>Басқару, әмір ету, боп-салау, өкім ету</i> и т.д.	ЕТ405
6. Глаголы подчинения	<i>Көну, бас иіу, ырқына көну</i> и т.д.	ЕТ406
<b>Семантический класс: V. Глаголы психической деятельности</b>		
1. <u>Глаголы ощущения:</u>	<i>Сезу, сезіну, байқау, абайлау, аңғару, пайымдау, түйсіну</i> ...	ЕТ501
а) глаголы ощущения в общем смысле		
б) глаголы видения	<i>Көру, көз салу, көз татау, үңілу,</i>	ЕТ502

	<i>бажырайу</i> и т.д.	
в) глаголы вкусовые	<i>Кермектену, ашқылтымдану, тату</i>	ET503
г) глаголы осязания (связанные с запахом)	<i>Бықсу, мүңку, аңқу, сасу, борсу, қаңсу</i> и т.д.	ET504
д) глаголы осязания, связанные с порами тела человека и животных	<i>Сыздау, ашу, қышу, тызылдау, сырқырау</i> и т.д.	ET505
е) глаголы, связанные со слухом	<i>Тыңдау, есіту, құлақ салу, дауыстау, естілу</i> и т.д.	ET506
2. <u>Глаголы мышления:</u>		
а) глаголы мышления, связанные с процессом мышления вообще	<i>Ойлау, пайымдау, ой жүгірту, ақылға салу, болжау, шамалау</i> и т.д.	ET507
б) глаголы предположения (приблизительности)	<i>Шамалау, болжау, тұспалдау, межеелеу</i> и т.д.	ET508
в) глаголы, связанные с выводами, заключениями, обобщениями	<i>Қорытындылау, шешілу, кесіп айту</i> и т.д.	ET509
г) глаголы доверия	<i>Сену, достасу, дос көру, жақындасу, сенім арту.</i>	ET510
д) глаголы запоминания	<i>Ескер, ұмытпа, есте сақта, зердеде қалдыру ...</i>	ET511
3. <u>Глаголы психического воздействия</u>	<i>Қуанту, күлдіру, қорқыту, ойландыру</i> и т.д.	ET512
4. <u>Глаголы внимания и желания</u>	<i>(жақсы, ақылды, сыпайы, ұстамды)болу</i> и т.д.	ET513
5 <u>Глаголы активной психической деятельности</u>	<i>Айқайлау, ұрысу, зар жылау, қарқылдау</i> и т.д.	ET514

6. Глаголы эмоциональные:		
а) глаголы восхищения	<i>Таңдану, қуану, таңырқау</i> и т.д.	ET515
б) глаголы радости	<i>Шаттану, қуану, мәз болу ...</i>	ET516
в) глаголы удовлетворения	<i>Рахаттану, қанағаттану, шаттану</i> и т.д.	ET517
г) глаголы эмоционального переживания (скорби)	<i>Қапалану, қайғыру, қабырғасы қайысу</i> и т.д.	ET518
<b>Семантический класс: VI. Глаголы физического состояния и нормы</b>		
1. Специфичные для элементов неживой природы (изменение цвета)	<i>Жылтырау, жарқырау, ағару, қарайу, сарғайу, сәулелену</i> и т.д.	ET601
2. Глаголы физического состояния	<i>Сақталу, бұзылмау, ашымау...</i>	ET602
а) обозначающие внутренние психологические состояния человека	<i>Қасарысу, ашулану, қайғыру, толқу, қуану, ойлану, армандау ...</i>	ET603
б) обозначающие субъективные черты человека (наружные)	<i>Кісімсу, адамсу, баласыну, сасқалақтау</i> и т.д.	ET604
3. Глаголы, передающие изменение объема предмета, вещества	<i>Шөгу, ұзару, солу, толысу, зораяу, кеңею</i> и т.д.	ET605
4. Глаголы, передающие изменение пред-	<i>Азаю, көбею, молаю, ортаю,</i>	

мета в количественном отношении	<i>қосылу</i> и т.д.	ЕТ606
5. Глаголы, обозначающие изменение внутренних свойств человека или предмета	<i>Іру, шіру, үгілу, тозу, босау, өзгеру</i> и т.д.	ЕТ607
6. Глаголы, употребляющиеся в связи с привычками людей	<i>Әдеттену, үйрену, дағдылану, жалқаулану ...</i>	ЕТ608
7. Глаголы, связанные с процессами рождения и роста, свойственными человеку или животным	<i>Туу, өсу, өну, ержету, жігіт болу</i> и т.д.	ЕТ609
<b>Семантический класс: VII. Глаголы психического состояния</b>		
1. Глаголы, свойственные только человеку	<i>Қайғыру, күңірену, ренжу, күлу, жылау, қуану</i>	ЕТ701
2. Глаголы, связанные с природными явлениями	<i>Азынау, гүлдеу, борандату, желдету</i> и т.д.	ЕТ702
3. Глаголы, свойственные человеку и особям животного мира	<i>Қайғыру, қуану, ашулану, ызалану</i> и т.д.	ЕТ703
4. Глаголы, приписываемые элементам неживой природы	<i>Қайғыру, мұңаю, күңірену, қуану, жадырау...</i>	ЕТ704
5. Глаголы, свойственные человеку, особям животного мира и приписываемые элементам неживой природы	<i>Қайғыру, мұңаю, жадырау, күңірену, қуану, күлімсіреу</i> и т.д.	ЕТ705
<b>Семантический класс: VIII. Фазисные и модальные глаголы</b>		
1. Начало действия	<i>Бастау, кірісу, жасағалы-, кіріскелі-</i>	ЕТ801
2. Продолжения действия	<i>Жалғастыру, созу, атқару, -ып, -іп (отыр)...</i>	ЕТ802
3. Конец действия	<i>Аяқтау, қою, бітіру, тоқтату ...</i>	ЕТ803
и т.д.	и т.д.	и т.д.

Далее мы решили опустить подробный табличный вид каузативных глаголов. В нашей таблице они заимствованы из рукописи кандидатской диссертации Г.М.Алиной «Каузальные отношения в системе русского и казахского языков» и делятся на следующие классы и подклассы:

I. Собственно-каузативные глаголы:

- 1) понуждение; 2) вызывание, причинение;
- 3) побуждение; 4) обуславливание; 5) воздействие;
- 6) допущение.

II. Частно-каузативные глаголы:

- 1) созидания объектов;
- 2) уничтожения объектов;
- 3) со значением рождения;
- 4) со значением лишения жизни кого-либо;
- 5) со значением изменения объектов;
- 6) со значением каузации контакта предметов;
- 7) со значением каузации функциональных отношений;

- 8) со значением каузации движения;
- 9) со значением каузации посесивных отношений;

Подводя итоги, можно сказать, что в национальном корпусе казахского языка имеются определённые трудности в разработке лексико-семантической разметки. Но, несмотря на то, что различать значение слова в языке, познание его смысла не составит трудностей для лингвистов, простому потребителю не легко справиться с семантической стороны слова. Важным делом в овладении языком является постановка лексико-семантических разметок к каждому слову в корпусе языка. В вышеприведенных таблицах важно различать, какое слово является слоганом, а определение к слову лингвистический анализ имеет важное значение при практическом подходе к корпусу.

Среди слов, входящих в морфологическую категорию, таких как существительное, глагол, прилагательное, и вместе с этим определить грамматические стороны искомого слова дает пользователю корпуса полноценные лексико-семантические знания. А лексико-семантическое развитие корпуса требует времени и теоретическое, практическое знание. При группировании семантической классификации существительных, глаголов, прилагательных имеют большое значение при разработке в Национальный корпус казахского языка лексико-семантической разметки

## СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. Алишер Навои (1991). «Арбаин хадис». Составители и комментарии Кабил Ахмад, Оллоберди Махмуд. Ташкент.
2. Алишер Навои (1991). Арбаин. – Ташкент, Мерос.
3. Арбаин хадис ва қирқ мақол (1994). Ташкент.
4. Алишер Навои (1991). “Арбаин хадис”. Ташкент.
5. Навои (2000). Мукамал асарлар тўплами. В16-и томах. Ташкент.
6. Каюмов, А. (2002). «Жемчужины мудрости». Ташкент.
7. Наджафова, А. (2002), Алишер Навои. «Арбаин» - «Сорок слов», Ташкентская правда –№ 32. Ташкент.
8. Arayana (1966). Vol.XXIV, 1-2, Jan. – Feb. –Kabul.

*Ayman Zhanabekova (Qazaxıstan)**Kunsulu Pirmanova (Qazaxıstan)**Sansızbay Madiev Sharapiuly (Qazaxıstan)**Bulbul Karbozova (Qazaxıstan)*

## **Qazax dilinin milli korpusunda leksik-semantik işarələrin hazırlanması**

### **Xülasə**

Məqalədə qazax dilinin milli korpusunun leksik və semantik işarələnməsinin hazırlanması üçün zəruri prinsiplər barədə ümumi məlumat verilir. Ümumiyyətlə, söhbət korpusun iki istiqamətindən gedir: Annotasiyalı və qeyri-annotasiyalı korpuslar. Korpus dilçiliyi qazax dilini öyrənənlərə, dilçilərə, linqvistlərə və korpusun digər istifadəçilərinə qazax dilinin nəzəri və praktik biliklərini mənimsəməyə kömək edəcəkdir.

Dünya dil korpuslarının nəzəriyyəsinə gəldikdə, korpusun daxili subkorpus funksiyası ilə bağlı kifayət qədər məlumatlar var. Dil korpusunda dil səviyyələrinin linqvistik təsnifatları mövcuddur. Dilin bütün səviyyələrinə aid praktik məsələlərin həlli leksik-semantik və digər işarələrin hazırlanması əsnasında çox vacib bir iş kimi ortaya çıxır. Bu məqalədə leksik-semantik işarələnmənin hazırlanması üçün zəruri praktiki məsələlər təsvir edilmişdir. Bu işarələnmə ilə bağlı korpusun algoritmini tərtib edərkən qazax dilçiliyində tətbiqi sahənin banisi professor A.K.Jubanovun hazırladığı qazax dilində dil hadisələrinin tezliyi modelinə əsaslandıq.

Qazax dilinin milli korpusunda leksik-semantik işarələnmənin hazırlanmasında bir sıra çətinliklər də var. Baxmayaraq ki, dilçilər üçün dildə sözün mənasını ayırd etmək, onun mənasını bilmək çətin olmayacaq, ancaq sırası bir istifadəçinin sözün semantik tərəfi ilə bağlı bu məsələnin öhdəsindən gəlməsi heç də asan iş deyil. Dilin mənimsənilməsində mühüm işlərdən biri dilin korpusunda hər bir söz üçün leksik və semantik işarələrin hazırlanmasıdır.

**Açar sözlər:** *korpus dilçiliyi, qazax dilinin milli korpusu, işarələmə, leksik-semantik işarələmə, qazax dili leksikasının semantik təsnifatı prinsipləri, isimlərin semantik təsnifatı, feillərin semantik təsnifatı.*



*Ayman Zhanabekova (Kazakhstan)*  
*Kunsulu Pirmanova (Kazakhstan)*  
*Sansyzbay Madiev Sharapiuly (Kazakhstan)*  
*Bulbul Karbozova (Kazakhstan)*

## **The Compiling of the Lexical-Semantic Labels in the National Corpus of Kazakh Language**

### **Abstract**

The general information about necessary principles for the compiling of lexical-semantic labeling of Kazakh National Corpus is given in the article. Generally, it is about the two directions of corpus. About the annotated and non-annotated corpora.

Corpus linguistics will help to the learners of Kazakh language, linguists and the other users of corpus for appropriating the theoretical and practical knowledges of Kazakh language. If we look through the theory of the world language corpora, we can see that there is enough information related with the function of internal sub-corpus of the corpus. The linguistic classifications of the language levels exist in the language corpus. The solving of the practical issues belonging to all levels of the language is very important work during the compiling of lexical-semantic and the other labels. The necessary practical issues have been described for the compiling of the lexical-semantic labeling in the article. During compiling the algorithm of corpus related with this labeling, we were based to the model of frequency of the cases of Kazakh language which was compiled by the founder of applied field in Kazakh linguistics, Professor A.K.Jubanov.

There are difficulties in the compiling of the lexical-semantic labeling in the National Corpus of Kazakh language, too. Though to make out the meaning of the word, to know its meaning won't be difficult for linguists in language, but to cope with this issue related with the semantic part of the word isn't easy for the simple language user. In the appropriation of the language, one of the important issues is the compiling of lexical-semantic labels for each word in the corpus of language.

**Key words:** *corpus linguistics, National Corpus of Kazakh language, labeling, lexical-semantic labeling, the principles of semantic classification of Kazakh language vocabulary, the semantic classification of nouns, semantic classification of verbs.*

---